International Journal of Technological Exploration and Learning (IJTEL)
ISSN: 2319-2135
Volume 2 Issue 5 (October 2013)

Human Activity Detection and Recognition Algorithm from Video Surveillances.

Kanchan Gaikwad
Department of Computer Engg.,
MGM College of Engineering
Kamothe, Panvel.

Manoj D. Patil
Department of Computer Engineering
Datta Meghe College of Engineering,
Airoli, Navi Mumbai

Abstract: **This paper presents a novel approach for automatic recognition of human activities from video sequences. Visual study of human motion is currently one of the most active research topics in computer vision. This strong interest is determined by a wide spectrum of promising applications in many areas such as virtual reality, smart surveillance, perceptual interface, etc. Human motion analysis concerns the detection, tracking and recognition of people. This paper includes the tracking of human activity from video sequencing images. Here we first make frames from video and apply GMM on it. Using HMM we classify the activity and detect the activity. The importance is on three major issues involved in a general human motion analysis system, namely human detection, tracking and activity understanding.**

Key words: Human activity recognition; HMM thresholding,; Gaussian filtering; Content Based Video Analysis ; Interactive Applications and Environments

I. INTRODUCTION

Active research topic in computer visions are the dynamic scenes detection, classifying object, tracking and recognizing activity and description of behavior. Visual surveillance strategies have long been in use together information and to monitor people, events and activities. Video surveillance works as to detect moving object [1], [2], [3], [4], classify [5], [6] the detected object track [7], them through the sequence of images and analysis the behaviors. Visual surveillance technologies [8], CCD cameras, thermal cameras and night vision device are the three most widely used devices in the visual surveillance market. The main goal of visual surveillance is not only to monitor, but also to automate the entire surveillance task. The goal of visual surveillance is to develop intelligent visual surveillance to replace the traditional passive video surveillance that is proving in effective as the numbers of cameras exceed the capability of human operators to monitor them. The automated surveillance systems can be implemented for both offline like storing the video sequence and to analyses the information in that sequence. But now days online surveillance system is very much needful in all public and private sectors due to predict and avoid unwanted movements, terrorist activities in those areas. It is helpful for traffic monitoring, transport networks, traffic flow analysis, understanding of human activity [11], [10], [9], home nursing, monitoring of endangered species, and observation of people and vehicles within a busy environment along many others to prevent theft and robbery. Some of the areas where video surveillance system place a major role in many application are 1) for military security 2)patrolling of country borders 3)extracting statistics for sport activities 4)surveillance of forests for fire detection 5) patrolling of highways and railway for accident detection.

Human Activity Recognition researches mostly observe human actions to obtain understanding on types of activities that humans perform within a time interval. Human Activity Recognition dominantly observes a series of physical actions that construct one physical activity. Physical activity is the essential human activity that mainly garners attention from many researchers to recognize.

The Proposed system gives human activity detection from online video surveillance and detects single human activity from video sequence. It is useful in many other applications. The paper is arranged as section1 is including introduction. The review of methods are include in section 2.The proposed system is describe in section 3. Advantages of proposed system are in section 4. The conclusion in section 5 and references are in section 6.

II. RELATED SURVEY FOR HUMAN ACTIVITY RECOGNITION

The importance and popularity of human activity analysis has led to several previous surveys. The earliest relevant review was probably due to Aggarwal et al. [12]. It covered various methods used in articulated and elastic non-rigid motion. As for articulated motion, the approaches with or without a prior shape models were described. Cedars and Shah [13] presented an overview of methods for motion extraction, in which human activity analysis was illustrated as action recognition, recognition of body parts and body configuration estimation.

Aggarwal and Cai gave review [15] covering 69 publications was an extension of their workshop paper [14]. The paper provided an overview of various tasks involved in motion analysis of human body. The focuses were on three major areas related to interpreting human motion: (a) motion analysis involving human body parts, (b) tracking moving human from a single view or multiple camera perspectives, and (c) recognizing human activities from image sequences.

A similar survey by Gavrila described the work in human activity analysis. Its highlighting on discussing various methodologies that were grouped into 2-D approaches with or without explicit shape models and 3-D approaches. It concluded with two main future directions in 3-D tracking and action recognition.

Recently, a significant study by Pentland [16] centered on person identification, surveillance/monitoring, 3-D methods, and smart rooms/perceptual user interfaces to evaluate the state of the art of "looking at people". The paper was not planned to survey the current work on human motion analysis, but touched on several interesting topics in human motion analysis and its applications.

The latest survey of computer-vision-based human motion capture was presented by Moeslund and Granum [17]. Its focus was on a general overview based on the taxonomy of system functionalities, viz. initialization, tracking, pose estimation and recognition. In addition, a number of general assumptions used in this research field were identified and suggestions for future research directions were offered.

III. HUMAN ACTIVITY RECOGNITION FROM VIDEO

In the proposed system live video is taking as input and then image segmentation is done by using Gaussian mixture model. The Gaussian mixture model has different methods for subtraction of background from image frame. Here we used the frame differencing method for selection of foreground image from sequences of frames. Further we used the Noise and shadow immune features of Gaussian mixture model. As shown in block diagram of proposed system as follows

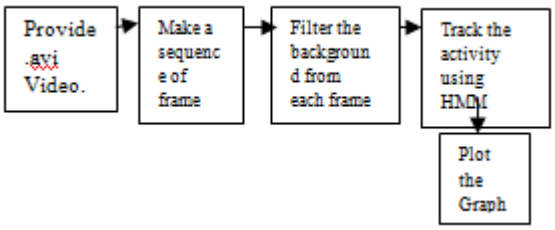

Figure 1. Shows the block diagram for human activity recognition.

This technique reduces the number of iteration for selecting human features from Sequences of frames. It means that we start by extracting the area that contains the person performing the activity, i.e., region of interest (ROI). Figure 2 shows an example result of background subtraction. A rectangular ROI is obtained from the result of background subtraction after noise and shadow removal as shown in Figure 2.b.

Figure 2. (a) A frame from a aerobic sequence, b) The ROI obtained after background subtraction.

The background subtraction equation becomes

$$B(x; y; t) = l(x; y; t - 1)$$
$$|l(x; y; t) - (x; y; t - 1)| > Th \quad (1)$$

$$K = argmin\{-\sum_{i=0}^{n} \ln f\left(xi\middle|\theta(K)\right) + \frac{K-1}{2}\ln(n) +$$

$$d2+ 3d4\ K \ln(n)\} \quad (2)$$

After we get the estimated model order, we can model each type of activity by Gaussian Mixture in the feature space. Then, the ROI is partitioned into 64 blocks, $B(k)$, with equal sizes, where $k = 1, \ldots, 64$. The average optical flow vector for every block is then computed by:

$$\overline{O_k} = \begin{bmatrix} \overline{O_{kx}} \\ \overline{O_{ky}} \end{bmatrix} = \frac{1}{n} \sum_{i,j \in B(k)} \begin{bmatrix} \bar{o}_x(i,j) \\ \bar{o}_y(i,j) \end{bmatrix} \quad (3)$$

where n is the number of pixels in a single block. Then, we compose the vector

$$O = \left[\bar{o}_1, \bar{o}_2 \ldots\ldots\ldots \bar{o}_{64}\right]^T$$

every frame to represent its motion feature vector, where each element contains two components for the x and the y directions.

As we get the foreground image now the selection of features like position and speed of object. The selection of position object is done by using the simple Average Theorem and combining velocity estimate over different time interval algorithm is used for selecting the speed of object.

Once the ROI (Region of Interest) is get then human activity is tracking from video using HMM thresholding algorithm. In this algorithm the maximum value of distance between frame sequences is considered as motion of object. We used both X and Y to denote video clips, and in the conditional probability expressions they represent the corresponding feature vector sequences. Based on this idea, the recognition result can be obtained as follows:

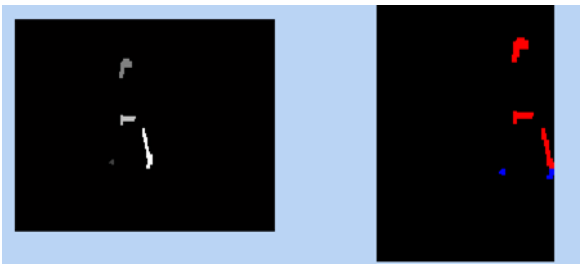

Figure 3. Tracking of activity from Region of Interest.

For recognition of activity we calculate the deviation. It is calculated for each matrix. In mat lab code The deviation is calculated as:

```
Deviation = sum (hsvCompare (pixels (:,:,:,f),motionMean,shadowLevel)./motion Deviation,3);
```

It considers the maximum distance between two frames as HSV Compare. Using the deviation the Sigma values is calculated in matlab as

```
sigma(:,:,1) = sigma(:,:,1)/(2*pi);
```

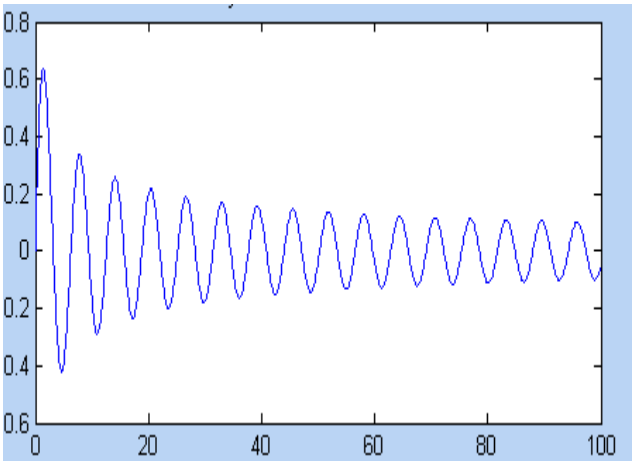

Figure 4. Graph of recognized activity position by speed of human activity.

IV. APPLICATION OF HAR

Increasing importance from researchers in human activity recognition field has various underpinnings. here are several probable applications that can be developed with human activity recognition enhancing the service significantly. Application examples can be application in smart homes, on-demand information systems, surveillance and monitoring systems, interactive interfaces for mobile services and games, up to healthcare application for both inpatient and outpatient treatment. Human activity recognition is mainly embattled towards expansion of intelligent healthcare system. Health problem is ultimately the critical issue that motivates researches to conduct researches in human activity recognition.

Researches for healthcare having the most significant attention is presented as in [18], [19], [20], and [21]. W. H. Wu et al. in their paper [18] propose MEDIC: a patient monitoring and medical diagnosis system architecture by using physiological body worn and wireless contextual sensors network. X. Long et al. in their paper [19] measure daily energy expenditure in daily activities and sport activities as a base on physical activity classification. Bartalesi et al. [20] interestingly use kinesthetic wearable sensor for stroke patients to detect their upper limb gesture where this works followed up in Tognetti et al. [21] that also interestingly use garment-based sensors like upper limb kinesthetic garment (ULKG) and sensing gloves.

Besides healthcare application, human activity recognition is also made use in smart homes application. Smart homes enable intelligent control by residential occupants to various automations in residences. Smart homes are equipped with various sensors to capture contexts in their surroundings and with intelligence to process the contexts so that they can respond flexibly according to the contexts.

Human activity recognition in smart homes mainly employs contexts' identification and recognition from multimodal features using multiple types of sensors. Seong Kim in [22] uses multimodal features that later being fused together. He obtains contexts from video sensor-based HAR system up to motion sensor-based HAR system. One interesting fact from his project named Smart homes: Personal Life Log is that he realized the potential of smart phone in human activity recognition in smart home perspectives. Therefore, he also has experienced using smart phone in his smart home project. This is interesting because smartphone is implemented in smart home application most dominantly as merely the remote controller. Liang-Wang et al. present their work in [23] which focus more in low-level activity features like human environment interaction in multimodal contexts using multiple wearable sensors.

V. ANALYSIS OF RESULT

The proposed method gives best result because; it uses the special feature of Gaussian mixture model to immune the noise and shadow from the image frame. It minimizes the no of iteration for selecting the foreground image and most promising features of object. The trained HMM track and recognized the activity of human. The proposed method uses the frames of video as shown in following figure 5 and tracked activity graph as shown in figure 6.

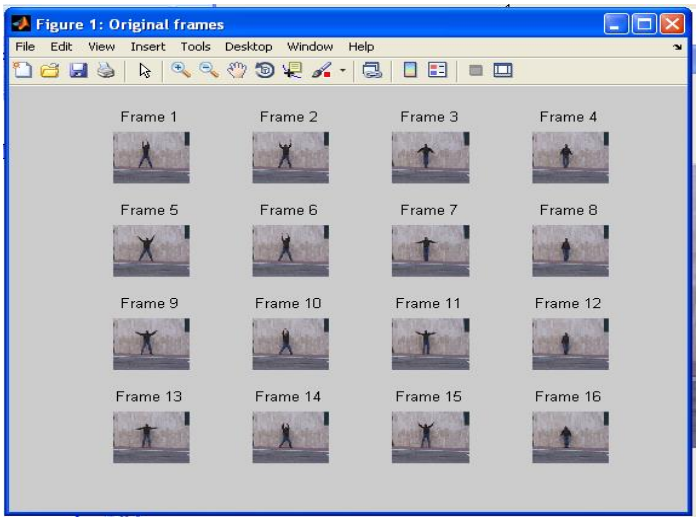

Figure 5. Shows the frames of video.

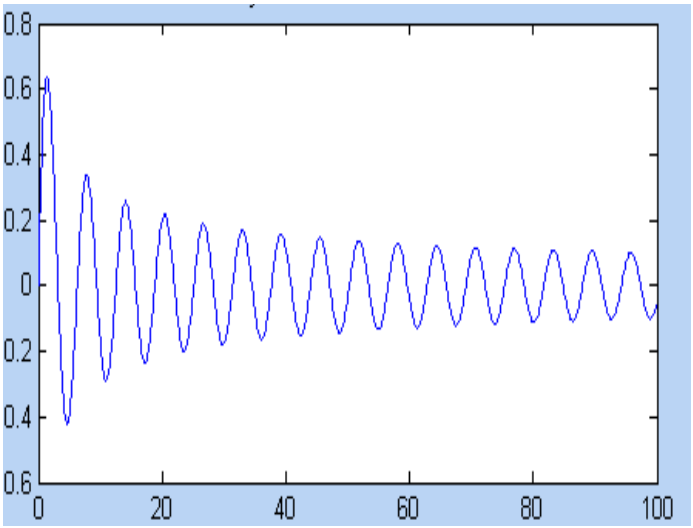

Figure 6. Graph of recognized activity position by speed of human activity.

The confusion matrix of this method is as follows:

TABLE I. CONFUSION MATRIX FOR ACTIVITY RECOGNITION IN(%)

Actions	Walking	Running	Gymnasium work	Standing
walking	98	1	-	1
Running	1	98	-	1
Gymnasium Work	1	1	97	1
Standing	1	1	-	98

TABLE II. COMPARISON BETWEEN PROPOSED METHOD WITH OTHER METHODS.

Sr.no	Method	Recognition Rate	Features considered
01	Human Activity Recognition Based on Spatial Transform in Video Surveillance	Walking- 88% Kick- 94% Sit down- 98%	Area, angle, centroid
02	HMM-based Human Action Recognition Using Multi view Image	Accuracy rate is 87%	silhouette, optical flow, and combined features
03	A Comparison of HMMs and Dynamic Bayesian Networks for Recognizing Office Activities	Using DBN 96.7% and using hmm 94.2%	Audio, Video, Keyboard and Mouse
04	Recognition of human activities using layered hidden markov models	Recognition rate 80%	Left and right hand movement
05	Proposed System	Walking 98% Running 98% Gymnasium work out 97% Standing 98% Over all recognition rate 98%	Position and Speed of object

The recognition rate of this method is good as compare with the other methods. This method require the less time for the recognition of activity of human. The error rate is reduces to 1.2%. The recognition rate of activity is increases and within less time. But this method is useful for the system which requires only the signals of recognized activity like instruction for robots. The comparison is shows as in Table 2.

VI. CONCLUSION

In this paper we use Gaussian mixture model for subtracting background from image frame sequence, it had the ability to remove noise from the image sequences and it helps to reduces the number of iteration to find activity from video. Once the background is get subtracted then the foreground image has to be select using frame differencing method for video. As we get the processed video the feature selection of image is done by Simple Average Theorem, combining velocity estimate over different time interval algorithm. Then using HMM we track and recognizes the motion of human from image sequence frames. Using this method the activity detection time is get reduces and error rate is also get reduces by 1.2%. The activity detection is very emerging topic in computer science world which is very useful to develop may applications like smart homes, on-demand information systems, health care application.

REFERENCES:

[1] M. Murshed, A.Ramirez, O.Chae, (2010) "Statistical Background Modelling: An Edge Segment Based Moving Object Detection Approach" IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS).

[2] Peng Chena, Xiang Chenb, Beibei Jina, Xiangbing Zhua (2012) "Online EM Algorithm for Background Subtraction".

[3] Lun Zhang, Li S.Z, Xiaotong Yuan, Shiming Xiang, (2007) "Real-time Object Classification in Video Surveillance Based on Appearance Learning" IEEE Conference on Computer Vision and Pattern Recognition.

[4] R.N Hota, V.Venkoparao, A.Rajagopal, (2007) "Shape Based Object Classification for Automated Video Surveillance with Feature Selection", 10th International Conference on Information technology, ICIT.

[5] Hong Lu, Hong Sheng Li, Lin Chai, Shu Min Fei, Guang Yun Liu, (2011) "Multi-Feature Fusion Based Object Detecting and Tracking", journal on Applied Mechanics and Materials.

[6] C. Beyan, A Temizel, (2012), "Adaptive mean-shift for automated multi object tracking" IET on Computer Vision.

[7] Blanco Adán, Carlos Roberto del and Jaureguizar Nuñez, Fernando and García Santos,(2011) "Bayesian Visual Surveillance, a Model for Detecting and Tracking a variable number of moving objects", IEEE International Conference on Image Processing, 13-14 de Enero del 2011.

[8] Arun Hampapyr, Lisa Brown, Jonathan Connell, Ahmet Ekin, Norman Haas,(2005) "Smart Video Surveillance Exploring the concept of multiscale spatiotemporal tracking", IEEE Signal Processing magazine, pp. 38-50.

[9] NeilM Robertson and IanD Reid, (2011) "Automatic Reasoning about Causal Events in Surveillance Video", EURASIP Journal on Image and Video Processing .

[10] Jenq-Neng Hwang, (2011) "Keynote talk 4: Automated understanding of video object events in a distributed smart camera network", International Conference on Advanced Technologies for Communications (ATC).

[11] Ali, S.F.; Jaffar, J.; Malik, A.S.,(2012) "Proposed framework of Intelligent Video Automatic Target Recognition System (IVATRs)", National Postgraduate Conference (NPC).

[12] J.K. Aggarwal, Q. Cai, W. Liao, B. Sabata, Articulated and elastic non-rigid motion: a review, Proceedings of the IEEE Workshop on Motion of Non-Rigid and Articulated Objects, 1994, pp. 2–14.

[13] C. Cedras, M. Shah, Motion-based recognition: a survey Image Vision Comput. 13 (2) (1995) 129–155.

[14] J.K. Aggarwal, Q. Cai, Human motion analysis: a review, Proceedings of the IEEE Workshop on Motion of Non-Rigid and Articulated Objects, 1997, pp. 90–102.

[15] J.K. Aggarwal, Q. Cai, Human motion analysis: a review, Comput. Vision Image Understanding 73 (3) (1999) 428–440.

[16] Alex Pentland, Looking at people: sensing for ubiquitous and wearable computing, IEEE Trans. Pattern Anal. Mach. Intell. 22 (1) (2000) 107–119..

[17] T.B. Moeslund, E. Granum, A survey of computer vision- based human motion capture, Comput. Vision Image Understanding 81 (3) (2001) 231–268.

[18] O. W. H. Wu, A. a T. Bui, M. a Batalin, L. K. Au, J. D. Binney, and W.J. Kaiser, "MEDIC: medical embedded device for individualized

care.," Artificial intelligence in medicine, vol. 42, no. 2, pp. 137-52, Feb. 2008.

[19] X. Long, B. Yin, and R. M. Aarts, "Single-Accelerometer-Based Daily Physical Activity Classification," pp. 6107-6110, 2009.

[20] R. Bartalesi, F. Lorussi, M. Tesconi, A. Tognetti, G. Zupone, and D. D. Rossi, "Wearable Kinesthetic System for Capturing and Classifying Upper Limb Gesture Kinesthetic Wearable Sensors and Kine- matic Models of Human Joints."

[21] A. Tognetti, F. Lorussi, M. Tesconi, R. Bartalesi, G. Zupone, and D. De Rossi, "Wearable kinesthetic systems for capturing and classifying body posture and gesture," 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference, pp. 1012-1015, 2005.

[22] T.-seong Kim, "Multi-modal Sensor-based Human Activity Recognition for Smart Homes Author biography," no. February, pp. 157-174, 2011.

[23] Y.-ting Chiang, Y.-ting Tsao, and J. Y.-jen Hsu, "A Framework for Activity Recognition in a Smart Home."